



PATENT OFFICE  
JAPANESE GOVERNMENT

RECEIVED  
JUL 22 2003  
Technology Center 2600

This is to certify that the annexed is a true copy of the following application  
as filed with this Office.

Date of Application : May 31, 1999

Application Number : P11-150855

Applicant(s) : NIPPON TELEGRAPH  
AND TELEPHONE CORPORATION

June 2, 2000

Commissioner,  
Patent Office                      Kouzou OIKAWA

Number of Certificate: 2000-3041001

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

N-105

RECEIVED

JUL 22 2003

Technology Center 2000

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

1999年 5月31日

出願番号

Application Number:

平成11年特許願第150855号

出願人

Applicant(s):

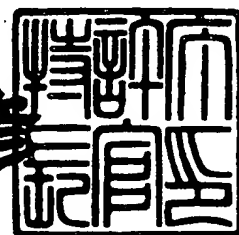
日本電信電話株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年 6月 2日

特許庁長官  
Commissioner,  
Patent Office

近藤 隆彦



【書類名】 特許願

【整理番号】 NTTH115181

【提出日】 平成11年 5月31日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 7/22

【発明者】

    【住所又は居所】 東京都新宿区西新宿三丁目 1 9 番 2 号 日本電信電話株式会社内

    【氏名】 大森 久美子

【発明者】

    【住所又は居所】 東京都新宿区西新宿三丁目 1 9 番 2 号 日本電信電話株式会社内

    【氏名】 東田 正信

【発明者】

    【住所又は居所】 東京都新宿区西新宿三丁目 1 9 番 2 号 日本電信電話株式会社内

    【氏名】 水澤 紀子

【特許出願人】

    【識別番号】 000004226

    【氏名又は名称】 日本電信電話株式会社

【代理人】

    【識別番号】 100066153

    【弁理士】

    【氏名又は名称】 草野 卓

【選任した代理人】

    【識別番号】 100100642

    【弁理士】

    【氏名又は名称】 稲垣 稔

【手数料の表示】

【予納台帳番号】 002897

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9806848

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 実時間音声対話制御方式を用いたデータベース検索方法、その装置及びプログラム記録媒体

【特許請求の範囲】

【請求項 1】 ユーザが音声入力する、検索対象音声検索キーをすべて含む音声認識用データベースに対して、音声対話のみを用いて音声検索キーの同定を行うデータベース検索方法において、

前記音声認識用データベースを階層構造化した階層別の複数の統計階層データベースとして保持し、下位の統計階層データベース程、構成語彙数を多くし、かつ最下位の統計階層データベースには全対象語彙を収容し、

すべての統計階層データベースに対して、並列に音声認識装置を用いて認識対象音声検索キーに対して音声認識処理を行い、音声検索キー候補と音声認識結果尤度（以下、検索キー認識尤度と呼ぶ）を 1 組のレコードとして持つ音声認識結果テーブルを各統計階層データベース毎に作成し、

各統計階層データベース中の規定尤度しきい値を越える検索キー認識尤度を持つ音声検索キー候補を有力候補とし、最上位の統計階層データベース中の有力候補が所定数以下かつ 0 でないか否かを調べる過程と、

所定数以下かつ 0 でないならば有力候補の中から候補の同定を行うための検索キー確定関連質問を、検索キー属性データベースを基に音声でユーザに提示する過程と、

検索キー確定関連質問に対するユーザの応答を、再度音声認識処理して音声検索キー候補同定の助けとなる関連属性情報として保持し、関連属性情報候補と認識結果尤度（以下、関連情報認識尤度と呼ぶ）を 1 組のレコードとして持つ関連情報認識結果テーブルを作成する過程と、

関連属性情報候補と、今絞込みの対象としている統計階層データベース中の音声検索キーの有力候補との関連性が、検索キー属性データベースを参照してあるかを調べ、関連性があれば、当該関連属性情報の関連情報認識尤度と当該音声検索キー有力候補の検索キー認識尤度とを統合する過程と、

音声検索キー有力候補の新たな認識尤度と定め、

最も大きな検索キー認識尤度を有する音声検索キー候補をユーザに提示し、正誤性を問うて音声検索キーの同定を行う過程と、

予め決めた次対話誘導条件を満たすかを調べる過程と、

次対話誘導条件を満たす場合はそれ以前に得られた関連属性情報候補が存在する場合は、当該関連属性情報候補との関連性を、存在しない場合は新たに検索キー確定関連質問を行い関連情報認識結果テーブルを作成し、獲得した関連属性情報候補との関連性を、次（下位）統計階層データベース中の音声検索キー候補について、検索キー属性データベースを基に判断して関連性が認められた検索キー候補に対して、音声検索キーの検索キー認識尤度と関連属性情報の関連情報認識尤度を統合し、今、絞込みの対象としている次階層統計階層データベース中の音声検索キー候補に対して、新たな認識尤度を定める過程と、

次階層統計階層データベース中の音声検索キー有力候補についてその数が所定数以下、かつ0でないか調べる前記過程を行い、それ以下の過程を行うことを繰り返し、音声検索キーの同定を行う過程とを有する

ことを特徴とする音声対話制御方式を用いたデータベース検索方法。

【請求項2】 請求項1記載のデータベース検索方法において、

前記複数の統計階層データベースは、

前記音声認識用データベースに対して、過去のデータベース検索装置に対するアクセス履歴などの統計的情報を用いて、データベース内の各データに対して重要度を定め、

最も重要度の高いデータから順に所定数のデータを選択して部分データベースを作成して統計階層データベースを作成し、

前記部分データベースを作成する際の所定数は、第 $n$ 階層に対して音声認識結果に基づいた音声検索キー同定のための対話がユーザとの間で行われている間に、第 $(n+1)$ 階層に対する音声認識及び検索キー認識尤度の算出、音声認識結果テーブルの作成が終了するように決定されている

ことを特徴とするデータベース検索方法。

【請求項3】 請求項1又は2記載のデータベース検索方法において、

前記関連情報認識尤度と前記検索キー認識尤度の前記統合は、

前記検索キー認識尤度に対して、音声検索キー候補として出力された全候補の検索キー認識尤度の合計が1になるように正規化し、

前記関連情報認識尤度に対しても全関連属性情報候補の認識尤度の合計が1になるように正規化し、

関連性が認められた当該音声検索キー候補の正規化された検索キー認識尤度と当該関連属性情報候補の正規化された関連情報認識尤度を乗算した結果を、新たな検索キー認識尤度とする

ことを特徴とするデータベース検索方法。

【請求項4】 請求項1乃至3の何れかに記載のデータベース検索方法において、

前記次対話誘導条件は

1. 上記音声検索キー有力候補が、上記所定数より多く出力された場合、
2. 上記音声検索キー有力候補が、存在しない場合、
3. 上記の検索キー確定関連質問から得られた関連属性情報候補との認識尤度の統合により、音声検索キーと同定して提示した候補が、ユーザから音声検索キーに該当しないと否定された場合、
4. 上記音声検索キー有力候補の中で、検索キー確定関連質問から得られた関連属性情報候補との関連性を持つ候補が、検索キー属性データベースを参照した結果存在しない場合、

の4つであることを特徴とするデータベース検索方法。

【請求項5】 入力音声を経電的音声信号に変換する音声入力手段と、

下位の階層データベース程、構成語彙数が多く、かつ最下位階層データベースは全対象語彙で構成された階層構造化した音声認識用データベースと、

上記音声信号を上記各階層データベースに対し、並列に音声認識する音声認識手段と、

音声検索キー候補と音声認識結果尤度（以下、検索キー認識尤度と呼ぶ）を1組のレコードとして持つ音声認識結果テーブルを各階層データベース毎に作成し、その作成終了順に出力する手段と、

上記音声信号を音声認識処理して、関連属性情報候補と認識結果尤度（以下、

関連情報認識尤度と呼ぶ)を1組のレコードとして持つ関連情報認識結果テーブルを作成して出力する手段と、

関連情報認識結果テーブルを保存する手段と、

音声検索キー候補と上記関連情報認識結果テーブル中の関連属性情報候補との関連性からその検索キー認識尤度と関連情報認識尤度とを統合して音声検索キー候補の新たな認識尤度とする手段と、

検索対象音声検索キーの全てについて各検索キー候補が有する属性項目を記録した音声検索キー属性データベースと、

音声検索キー有力候補が所定数より多い場合、或いは存在しない場合、或いはユーザに提示した音声検索キーが該当する候補でないと否定された場合、或いは関連属性情報候補と関連性のある候補が音声検索キー有力候補の中に存在しない場合、次階層データベースに検索対象を移動させる手段と、

上記音声検索キー有力候補が所定数より少なく、かつ0でない場合は上記検索キー属性データベースを参照して検索キー確定関連質問を決定する手段と、

音声認識結果テーブル中の検索キー認識尤度が最も大きい音声検索キーを提示することを決定する手段と、

上記決定された検索キー確定関連質問、上記決定された音声検索キー提示などに基づき質問文を作成する応答文/質問文作成手段と、

上記作成された質問文をユーザに対して音声出力する音声出力手段と、

を具備する実時間音声対話制御方式を用いたデータベース検索装置。

【請求項6】 入力された音声検索キーを受付ける処理と、

音声認識用データベースを階層構造化した階層別の複数の統計階層データベースに対して、並列に音声認識装置を用いて入力音声検索キーに対して音声認識処理を行い、音声検索キー候補と音声認識結果尤度(以下、検索キー認識尤度と呼ぶ)を1組のレコードとして持つ音声認識結果テーブルを各統計階層データベース毎に作成する処理と、

各統計階層データベース中の規定尤度しきい値を越える検索キー認識尤度を持つ音声検索キー候補を有力候補とし、最上位の統計階層データベース中の有力候補が所定数以下かつ0でないかを調べる処理と、



有力候補が所定数以下かつ0でなければ有力候補の中から候補の同定を行うための検索キー確定関連質問を、検索キー属性データベースを基に音声でユーザに提示する処理と、

検索キー確定関連質問に対するユーザの応答を、再度音声認識処理して音声検索キー候補同定の助けとなる関連属性情報として保持し、関連属性情報候補と認識結果尤度（以下、関連情報認識尤度と呼ぶ）を1組のレコードとして持つ関連情報認識結果テーブルを作成する処理と、

関連属性情報候補と、今絞込みの対象としている最上位統計階層データベース中の音声検索キーの有力候補との関連性を、検索キー属性データベースを参照することにより求め、当該関連属性情報の関連情報認識尤度と当該音声検索キー有力候補の検索キー認識尤度とを統合して、

音声検索キー有力候補の新たな認識尤度と定める処理と、

最も大きな検索キー認識尤度を有する音声検索キー候補をユーザに提示し、正誤性を問う処理と、

予め決めた次対話誘導条件を満たすか否かを調べる処理と、

次対話誘導条件を満たせば次階層（下位）統計階層データベース中の音声検索キー候補の中で、それ以前に得られた関連属性情報候補が存在する場合は、当該関連属性情報候補との関連性を、存在しない場合は新たに検索キー確定関連質問を行い関連情報認識結果テーブルを作成し、獲得した関連属性情報候補との関連性を、検索キー属性データベースを基に判断して関連性が認められた検索キー候補に対して、音声検索キーの検索キー認識尤度と関連属性情報の関連情報認識尤度を統合し、今、絞込みの対象としている次階層統計階層データベース中の音声検索キー候補に対して、新たな認識尤度を定める処理と、

次階層統計階層データベース中の音声検索キー有力候補数が所定数以下、かつ0でないかを調べる前記処理を行い、それ以下の処理を繰り返し音声検索キーの同定を行う処理と

をデータベース検索装置のコンピュータに実行させるプログラムを記録した記録媒体。

【請求項7】 請求項6記載の記録媒体において、

前記音声認識用データベースは、過去のデータベース検索装置に対するアクセス履歴などの統計的情報を用いて、データベース内の各データに対して重要度を定め、最も重要度の高いデータから順に所定数のデータを選択して部分データベースを作成し、

前記部分データベースを、下位階層ほど含有データ数が多く、最下位階層は前記音声認識用データベースの全データを含むように階層化され、かつ

前記部分データベースを作成する際の所定数は、第  $n$  階層に対して音声認識結果に基づいた音声検索キー同定のための対話がユーザとの間で行われている間に、第  $(n+1)$  階層に対する音声認識及び検索キー認識尤度の算出、音声認識結果テーブルの作成が終了するよう決定されていることを特徴とする記録媒体。

【請求項 8】 請求項 6 又は 7 記載の記録媒体において、

前記関連情報認識尤度と前記検索キー認識尤度の前記統合は、前記検索キー認識尤度に対して、音声検索キー候補として出力された全候補の検索キー認識尤度の合計が 1 になるように正規化する処理と、

前記関連情報認識尤度に対しても全関連属性情報候補の認識尤度の合計が 1 になるように正規化する処理と、

当該音声検索キー候補の正規化された検索キー認識尤度と当該関連属性情報候補の正規化された関連情報認識尤度を乗算した結果を、新たな検索キー認識尤度と処理とにより行う

ことを特徴とする記録媒体。

【請求項 9】 請求項 6 乃至 8 の何れかに記載の記録媒体において、

前記次対話誘導条件は

1. 上記音声検索キー有力候補が、上記所定数より多く出力された場合、
2. 上記音声検索キー有力候補が、存在しない場合、
3. 上記の検索キー確定関連質問から得られた関連属性情報候補との認識尤度の統合により、音声検索キーと同定して提示した候補が、ユーザから音声検索キーに該当しないと否定された場合、
4. 上記音声検索キー有力候補の中で、検索キー確定関連質問から得られた関連属性情報候補との関連性を持つ候補が、検索キー属性データベースを参照した

結果存在しない場合、

であることを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、与えられた一定時間内に処理不可能な検索対象語彙数を有する大規模検索データベースに対して、ユーザに対して音声認識処理及びデータベース照合時間を感じさせることなく、かつユーザとの間に音声認識処理の欠陥などによる不自然な対話を生じさせることなく、ユーザが入力した検索キーの同定を、自然な対話を通して実現する対話制御方法を採用した音声対話を用いたオペレータレス対話型検索手法、装置及びプログラム記録媒体に関するものである。

【0002】

【従来の技術】

従来の音声対話型システムは、ユーザが入力した音声に対して音声認識処理を行い、認識対象語彙をすべて含んでいるデータベースに対して検索して、音声認識処理及び認識結果に基づく候補の出力が定まるまでユーザを待機させ、音声認識処理及び認識候補出力が終了（決定）した時点で、音声認識尤度の高い順に候補をユーザに提示し、正しい検索キーが同定できるまで候補の提示を繰り返す。

【0003】

【発明が解決しようとする課題】

機械による音声認識処理では、特に不特定話者の音声、話速が一定でない発話に対する音声は認識が困難であり、さらに周囲の雑音などによる音質の劣化が原因となり、100%の音声認識精度が得られないことから、ユーザが音声により入力した音声検索キーを音声認識処理のみを用いて瞬時に同定することは困難である。

【0004】

また、音声対話型システムにおいて、ユーザとの間に自然な対話を実現する際、ユーザの入力に対してシステムは、人間が不自然さを感じない実時間に応答を返すことが大前提となる。しかし、与えられた一定時間内に音声認識処理可能な

語彙数には限界がある。そのため、一定時間内に処理不可能な語彙数を有する大規模データベースを認識対象とする場合、システムが音声認識処理の際にデータベース検索に要する処理時間をユーザに感じさせず、かつシステムの音声認識精度の欠陥をも感じさせずに、ユーザとシステムの自然な対話の中で一定時間にユーザが要求するタスクを達成することは困難であり、ユーザはシステムが認識処理結果を出力するまでの間、待機する必要がある、さらには提示された結果が誤認識の場合、再度認識結果が提示されるまで待機しなければならないことになり、音声を入力インタフェースとする、オペレータ対応と同等の迅速さ、正確さを有するシステムの構築は、現在の音声認識技術では困難とされている。

【0005】

この発明は、ユーザが入力する音声検索キーの同定をタスクとするオペレータレス音声対話型検索システムにおいて、ユーザに待機時間を感じさせずに、オペレータ対応のシステム同様の迅速性及び自然性を兼ね備えた対話を通してタスク達成を行う方法、装置及びプログラム記録媒体を提供することを目的とする。

【0006】

【課題を解決するための手段】

この発明は、ユーザとシステム間のタスク達成のために行われる対話応答に自然さを保つために、与えられた一定時間内に処理不可能な大規模検索対象語彙数を有する音声認識用データベースに対して、過去のアクセス頻度や使用頻度などの統計的情報により、データベース中の各データに対して重要度を定め、最も重要度の高いデータから順に所定数のデータを選択して部分データベースを作成し、部分データベースを、下位階層ほど含有データ数が多く、最下位階層は前記音声認識用データベースの全データを含むように階層構造化した複数の統計階層データベースとして保持する。

【0007】

実時間性を仮想的に実現するために、各階層のデータ数の相違による処理時間の差異を利用して、統計階層データベースの各階層に対し、並列に音声認識処理及び音声認識尤度に基づく音声検索キー候補の抽出を行い、処理が早く終了すると思われる最上位の統計階層データベースから得られる結果から順に、ユーザと

の対話誘導を行う。

【0008】

さらに、この発明によるタスク達成のためにユーザとの間で行われる対話制御方式は、図14に示すように、

まず始めに、最上位の統計階層データベースから得られる音声認識結果テーブルに対して（S1）、規定尤度しきい値以上の検索キー認識尤度を持つ音声検索キー有力候補が、所定数、例えば2よりも少数かつ0でない場合は（S2）、

音声検索キー絞込み方式に従い、関連属性情報獲得のための検索キー確定関連質問をユーザに対して行い（S3）、検索キー確定関連質問に対するユーザの応答を、再び音声認識装置を用いて音声認識処理して関連情報認識結果テーブルを作成し（S4）、得られた関連属性情報候補と、今絞込みの対象としている最上位統計階層データベース中の音声検索キー有力候補との関連性が、検索キー属性データベースを参照することにより認められた場合（S5）、当該関連属性情報の関連情報認識尤度と当該音声検索キー有力候補の検索キー認識尤度とを統合し、音声検索キー有力候補の新たな認識尤度と定め（S6）、

最も大きな検索キー認識尤度を有する音声検索キー候補をユーザに提示し、正誤性を問う（S7）。

【0009】

次対話誘導条件としては例えば以下の場合を予め決めておく。

1. 音声検索キー有力候補が、所定数より多く出力された場合、
2. 音声検索キー有力候補が、存在しない場合、
3. 検索キー確定関連質問から得られた関連属性情報候補との認識尤度の統合により、音声検索キーと同定して提示した候補が、ユーザから音声検索キーに該当しないと否定された場合、
4. 音声検索キー有力候補の中で、検索キー確定関連質問から得られた関連属性情報候補との関連性を持つ候補が、検索キー属性データベースを参照した結果存在しない場合、

前記音声検索キー有力候補が所定数よりも少数かつ0でない場合以外であって、前記4つの次対話誘導条件のいずれかを満たす場合（S8）、この時点までに

音声認識処理及び音声認識結果テーブル作成が終了している次階層（下位）統計階層データベース中の（S12）、音声検索キー有力候補の中で、既に得られた関連属性情報候補が存在する場合は（S9）、関連属性情報候補との関連性、存在しない場合は、新たに検索キー確定関連質問を行い（S10）、関連情報認識結果テーブルを作成して（S11）、得られた関連属性情報候補との関連性が、検索キー属性データベースを参照することにより認められた場合（S13）、音声検索キー有力候補に対して、検索キー認識尤度と関連属性情報の関連情報認識尤度を統合し、新たな検索キー認識尤度を定め（S14）、

認識対象の次階層統計階層データベース中の音声検索キー有力候補数が、再び所定数以下、かつ0でないならば（S2）、

前記音声検索キー候補絞り込み方式に従って、他の関連属性情報についての検索キー確定関連質問をユーザに提示し（S3）、得られた関連属性情報候補だけでなくそれ以前の検索キー確定関連質問から得られているすべての関連属性情報候補との関連性を総合的に判断し（S5）、今認識対象としている次階層統計階層データベースの音声検索キー有力候補に対して、関連性のある関連属性情報候補との認識尤度を統合し（S6）、最も大きい検索キー認識尤度を持つ音声検索キーを候補と同定してユーザに提示し正誤性を問う（S7）。

【0010】

さらに、次階層統計階層データベースに対する音声認識の結果が、前記4つの次対話誘導条件のいずれかを満たす場合（S8）、

その時点で音声認識処理及び音声認識結果テーブル作成が終了している、さらに下位の次階層統計階層データベースに対して、上位の統計階層データベースに対する処理同様に（S9, S10, S11, S12）、それまでに得られているすべての関連属性情報候補との関連性が認められた場合は認識尤度を統合して、新たな検索キー認識尤度を算出し（S14）、

音声検索キー有力候補数が所定数以下かつ0でないならば（S2）、音声検索キーが同定できるまで、検索キー確定関連質問を行い、音声検索キー絞り込み方式を適用し、

次対話誘導条件を満たすならば、さらに次階層統計階層データベースに対して

、それまでに得られているすべての関連属性情報候補との認識尤度の統合を繰り返し、音声検索キーの同定を行う

ことを特徴とする。

【0011】

この発明において使用する統計階層データベースは、この発明において、音声検索キーの絞込みを行うために音声検索キーが有する関連属性情報を尋ねる検索キー確定関連質問を行う際、検索対象である音声認識用データベースの全データに対して、各データの特徴を表す属性項目の属性値を、関連属性情報として保持している検索キー属性データベースを使用する

ことを特徴とする。

【0012】

またこの発明において、音声検索キーの有力候補から候補の絞込みを行うために、検索キー確定関連質問から得られる複数の関連属性情報候補と、絞込みの対象としている音声検索キー有力候補との間の関連性が、検索キー属性データベースを参照することにより認められた場合、検索キー認識尤度と関連情報認識尤度を正規化及び乗算することで、候補の同定を実現する

ことを特徴とする。

作用

音声検索キーの同定を目的とするデータベース検索装置において、音声検索キー同定のためにユーザとの間で行われる対話応答に自然さを保つ範囲の一定時間内に音声認識処理及びデータベース照合不可能な大規模音声認識対象語彙数を有する検索データベースを検索対象とする場合、この発明は、オペレータ対応のシステム同様ユーザに音声認識処理及びデータベース照合時間、機械による音声認識精度の欠陥を感じさせることなく、さらにユーザとの対話において、自然性を重視した対話制御方式を用いて音声検索キーの同定を実現する。

【0013】

この発明のデータベース検索方法では、検索対象データベースが大規模であるため、検索データベースを重要度の高い順に階層構造化して複数の統計階層データベースとして保持し、第  $n$  階層に対する認識結果を基にした音声検索キー同定

のための対話がユーザとの間で行われている間に、第( $n+1$ )階層に対する音声認識及び検索キー認識尤度の算出、音声認識結果テーブルの作成が終了するよう各統計階層のデータ含有数を構成し、各統計階層のデータ数の相違による処理時間の差異を利用することで、ユーザに不自然さを感じない一定時間内の、音声検索キーに対する音声認識処理及び認識候補の出力を仮想的に実現する。

## 【0014】

そして、音声認識精度の欠陥をユーザに感じさせず、あたかも人間対人間の対話のような対話を対システム間においても成立させるために、前記統計階層データベースの各階層に対して並列に音声認識処理を行い、各階層毎に音声検索キー候補の抽出を行い、データ数が最小かつ統計的に重要度の高い音声検索キー候補を含む最上位の統計階層データベースに対する音声認識処理が最も早く終了することを利用して、最上位の統計階層データベースの音声認識結果テーブルから順に参照し、規定尤度しきい値を越える音声検索キー有力候補数に応じて、ユーザに対する対話誘導方法を決定する。

## 【0015】

音声検索キー有力候補数が、所定数以下かつ0でない場合は、有力候補の中から候補を絞込むための検索キー確定関連質問を行い、得られる関連属性情報候補との関連性が認められた、音声検索キー有力候補を音声検索キーに同定しユーザに提示する。

上記音声検索キー有力候補が所定数よりも多い場合、或いは存在しない場合、或いは上記のユーザに提示した音声検索キーが該当する候補ではないと、ユーザから否定された場合、或いは、上記検索キー確定関連質問により得られた関連属性情報候補と関連性のある候補が音声検索キー有力候補の中に存在しなかった場合、該当音声検索キーが最上位統計階層データベース中に存在しない可能性があることから、この時点で音声認識処理が終了している次階層(下位)統計階層データベースに検索対象を移動する。ユーザには、検索対象データベースが下位に移動したことを感じさせない。

## 【0016】

検索対象統計階層データベースを下位に移動した時点で、次対話誘導条件3、



4 の場合は、既に検索キー確定関連質問で得られている関連属性情報候補を再び利用して、次対話誘導条件 1、2 の場合は新たに検索キー確定関連質問を行い、得られた関連属性情報を利用して、今認識対象としてなっている第二統計階層データベース中の音声検索キー候補の中で、関連属性情報候補との関連性が認められた候補に対しては認識尤度の統合を行い、新たな認識尤度を定め、

再び、音声検索キー有力候補数を見て、有力候補が所定数以下かつ 0 でないならば、再び別の関連属性情報を尋ねる検索キー確定関連質問をして、新たに得られる関連属性情報候補も併せて利用して、第二統計階層データベース中の音声検索キー有力候補の絞り込みを行い、最上位統計階層データベースと同様に、認識尤度を統合した結果、最も大きな検索キー認識尤度を持つ音声検索キーをユーザに提示する。

#### 【0017】

認識対象である二番目の統計階層データベースの認識結果が、次対話誘導条件のいずれかを満たす場合は、最上位統計階層データベースの場合と同様に、第三統計階層データベースへと検索対象を移動し同様の対話誘導を、音声検索キーが同定できるまで繰り返す。

各統計階層データベースにおいて、音声検索キー有力候補が所定数以下かつ 0 でない場合の対話誘導においては、検索キー確定関連質問を行うことにより有力候補の持つ検索キー認識尤度の信頼性を高め、効果的に候補を絞り込むことができ

検索対象データベースを下位階層へ移動する対話誘導においては、下位階層ほど音声認識対象語彙数が多く認識精度の悪化が考えられるが、検索対象を下位に移動する時点までに獲得しているすべての関連属性情報候補との関連性を考慮することにより、より多くの情報を組み合わせて絞り込みを行うことでデータ数増加による認識精度の悪化を補うことができる。

#### 【0018】

また、この発明のデータベース検索方法は、音声認識精度が 100% でないことから音声検索キーを 100% の精度で同定することが不可能であるので、音声検索キーの持つ関連属性情報を用いて該当する音声検索キーの同定を試みる。し

かし、関連属性情報も検索キー確定関連質問の応答に対して音声認識を行うことで獲得するので、100%の精度で得られるわけではない。

【0019】

そこで考えられる音声検索キー候補と関連属性情報候補の認識尤度を正規化、統合することで音声認識精度の欠陥を補い、検索キー確定関連質問を行い候補の絞り込みを行いながら自然性を重視した対話制御方式を用いることで、ユーザには音声認識精度の欠陥を感じさせないように音声検索キー候補の絞り込みを行う。

音声認識用データベースの階層構造化、音声認識尤度の正規化、統合を利用した対話制御方式に基づいて、ユーザとの対話を行うことにより、高精度な自然性を兼ね備えた、ユーザに待機時間及び認識精度の欠陥を感じさせない、オペレータ対応しているようなデータベース検索装置の実現が可能になる。

【0020】

【発明の実施の形態】

図1は、この発明によるデータベース検索装置の実施形態の一例を示すブロック図である。このデータベース検索装置1は、音声入力部2、音声識別部3、対話制御部4、音声検索キー関連性計算部5、音声出力部6を備えている。音声識別部3は、音声認識部3-1、音声認識結果出力部3-2から構成される。対話制御部4は結果調整部4-1、対話誘導部4-2、応答文／質問文作成部4-3から構成される。音声識別部3では音声認識装置8を、音声出力部6では音声出力装置9をそれぞれ使用する。また、音声識別部3における入力音声に対する音声認識、及び対話制御部4の対話誘導部4-2における次の対話誘導の際、音声認識用データベース7を使用する。音声認識用データベース7は、複数の統計階層データベース7-1、検索対象音声検索キーすべてについて各検索キー候補が有する属性項目を記した音声検索キー属性データベース7-2、関連情報認識結果テーブル保存場所7-3、Yes／Noの認識データベース7-4から構成される。

【0021】

図2に統計階層データベース7-1の全体像の一例を示す。統計階層データベース7-1は、音声認識対象語彙であるすべての音声検索キー候補に対して、過

去のシステム利用者からのアクセス頻度などの統計的情報に基づいて重要度を定め、重要度の順に細分化及び階層構造化することにより作成される。音声入力部 2 は、ユーザから入力された音声音声識別部 3 に対して入力する。

【0022】

音声識別部 3 では、まず始めに音声認識部 3-1 が、音声入力部 2 により入力された音声に対して音声認識装置 8 を用いて音声認識処理を行う。その際、音声認識装置 8 は、入力された音声対話誘導のどの場面のものなのかに応じて音声認識用データベース 7 を参照する。すなわち、

ユーザから音声検索キーが入力されると同時に、統計階層データベース 7-1 の各階層に対して並列に音声認識処理プロセスが走り始める。

【0023】

・音声入力部 2 から、検索キー確定関連質問に対する応答が入力された場合は検索キー属性データベース 7-2、及び関連情報認識結果テーブル保存場所 7-3 を参照し、

・音声検索キー候補の提示に対する応答が入力された場合は、Yes/No の認識データベース 7-4、  
を参照する。

【0024】

そして、音声認識結果出力部 3-2 は、統計階層データベース 7-1 の各階層に対して複数並列に走っている音声認識処理プロセスが終了した順に、各階層毎に検索キー候補を、認識尤度を大きい順に並べた音声認識結果テーブルを作成する。図 3 に最上位統計階層データベースに対する音声認識結果テーブルの一例を示す。

【0025】

認識対象語彙数の相違から、統計階層データベース 7-1 中の最上位階層データベースに対する音声認識処理及び音声認識結果テーブル作成が最も早く終了する。最上位の統計階層データベースの音声認識結果テーブルが作成された時点で、対話制御部 4 へ認識結果を送る。この時、下位の統計階層データベースに対する認識処理、及び音声認識結果テーブル作成のプロセスは他の処理が次へ進んで

も続けて行う。

【0026】

対話制御部4では、音声識別部3の音声認識結果出力部3-2から送られる最上位統計階層データベースの音声認識結果テーブル中の規定尤度しきい値を超える検索キー認識尤度を持つ音声検索キー有力候補数に応じて、データベース検索装置1が、ユーザに対して次に行うべき対話誘導を決定する。

・結果調整部4-1に、音声検索キーに対する音声認識結果テーブルが送られてきた場合、

ー認識対象統計階層データベース中の音声検索キー有力候補がデータベース検索装置1による所定数よりも少なく、かつ0でない場合は、対話制御部4-2によって検索キー確定関連質問を検索キー属性データベース7-2（図4参照）を参照して行うことを決定し、応答文／質問文作成部4-3によって、検索キー確定関連質問の作成を行う。次対話誘導条件としては例えば以下の場合を予め決めておく。

【0027】

1. 音声検索キー有力候補が、所定数より多く出力された場合、
2. 音声検索キー有力候補が、存在しない場合、
3. 検索キー確定関連質問から得られた関連属性情報候補との認識尤度の統合により、音声検索キーと同定して提示した候補が、ユーザから音声検索キーに該当しないと否定された場合、
4. 音声検索キー有力候補の中で、検索キー確定関連質問から得られた関連属性情報候補との関連性を持つ候補が、検索キー属性データベースを参照した結果存在しない場合。

【0028】

ー認識対象統計階層データベース中の認識結果が前記次対話誘導条件のいずれかに当てはまる場合、対話制御部4-2は、認識対象が最上位の場合に限り関連属性情報が1つも得られていない場合は新たに検索キー確定関連質問を行うことを決定し質問文の作成を応答文／質問文作成部4-3に支持する。その他の場合は関連情報認識結果テーブル保存場所7-3を参照して、それまでに得られてい

る関連属性情報候補と認識対象統計階層データベース中の音声検索キー候補との関連性を、検索キー属性データベース 7-2 を参照して判断し認識尤度の正規化及び統合を行う（音声検索キー関連性計算部 5 にアクセス）。その結果新たに計算された検索キー認識尤度が最も大きい音声検索キーを提示する質問文の作成を応答文／質問文作成部 4-3 へ指示する。

## 【0029】

この間も、音声識別部 3 の音声認識部 3-1 及び音声認識結果出力部 3-2 によって、統計階層データベース 7-1 の各統計階層に対する音声認識、及び音声認識結果テーブル作成プロセスは動いている。

そして、作成したそれぞれの応答文／質問文を、ユーザに対して音声出力装置 9 を用いて音声出力部 6 から出力し、ユーザの応答を再び音声入力部 2 から得る。音声識別部 3 は、音声入力部 2 から入力された応答文／質問文に対するユーザの応答に対して、再び音声認識処理を行い結果を出力する。

## 【0030】

この時点で、第二統計階層データベースに対する音声認識結果テーブルの作成は終了している。

音声認識結果出力部 3-2 は、音声入力部 2 から送られてくるユーザの応答が検索キー確定関連質問への応答である場合、音声認識部 3-1 で音声認識処理をした結果から関連情報認識結果テーブルを作成し、音声認識用データベース 7 の関連情報認識結果テーブル保存場所 7-3 に保存をすると同時に結果調整部 4-1 へ結果を送る。図 5 に関連情報認識結果テーブルの一例を示す。

## 【0031】

・結果調整部 4-1 に、関連情報認識結果テーブルが送られた場合は、対話誘導部 4-2 は、既に音声認識及び音声認識結果テーブル作成が終了している第二統計階層データベースの音声認識結果テーブルを参照して、最上位の統計階層データベースに対する対話誘導同様規定尤度しきい値を超える音声検索キー有力候補数に応じて対話の方針を定める。

## 【0032】

第二統計階層データベースの音声認識結果テーブル中の音声検索キー有力候補

数が所定数以下かつ0でない場合は、検索キー確定関連質問による絞込みを行い、次対話誘導条件のいずれかに当てはまる場合は、それまでに得られている関連属性情報候補との関連性を判断して、認識尤度の統合を行い最も大きな検索キー認識尤度を算出した音声検索キー候補を音声検索キーと同定する。

#### 【0033】

・結果調整部4-1に、音声検索キーの提示に対する応答が送られた場合、応答が「Yes」の場合は、音声検索キーが確定できたことを知らせる応答文を作成することを対話誘導部4-2は決定し、応答文／質問文作成部4-3において応答文作成を行い処理を終了する。逆に、結果調整部4-1が「No」という応答を得た場合、次対話誘導条件を満たすことから、結果調整部4-1は対話誘導部4-2に対してさらなる対話誘導を指示し第三統計階層データベースの認識結果を用いた対話誘導が始まる。

#### 【0034】

このようにして音声検索キーが同定できるまで、音声検索キー有力候補数に応じた対話方針に従い、検索キー確定関連質問に得られる関連属性情報を利用した認識尤度の正規化、統合を繰り返す。

#### 実施例

以下、具体例を用いて説明する。

#### 【0035】

図1を用いてこの発明のデータベース検索方法により、コンサートチケット予約をタスクとする「チケット予約システム」のユーザが入力したチケット名の同定を行う流れを説明する。

チケット予約システムにおいて、システムが規定する尤度しきい値を3500、ユーザが入力した音声検索キーに対する音声認識処理の出力結果に対して、規定尤度しきい値3500を越える認識尤度を持つ検索キー候補がいくつ以下の場合にユーザに対して検索キー確定関連質問を行い、いくつ以上の場合に認識対象データベースを次階層へと移すかという、システムが対話誘導のために規定する、規定尤度しきい値以上の有力候補所定数を2とする。

#### 【0036】

以下、ユーザが「group」のコンサートチケットを予約する場合の流れを説明する。コンサートチケット予約システムにおいて、検索データベース全体としてコンサート主催者名350名をデータとして有する。このコンサート主催者350名をアクセス頻度（実施例の場合、過去1年間のCD売上などを基に発表されている人気ランキングを利用）を基に、四階層の統計階層データベースに分割する。図6に示すように、第一階層（最上位階層）は最も人気があるとされる上位60名のリスト、第二階層は第一階層のリスト60名に人気61位から150位までの90名を加えた150名のリスト、第三階層はさらに150位以下100名を加えた250名のリスト、第四階層はさらに250以下100名を加えた350名全員のリストから構成される。対象となる音声検索キー「group」は人気ランキング90位であるため最上位の第一統計階層データベースには存在しない。

#### 【0037】

音声入力部2から「group」という音声検索キーが入力されると、音声識別部3の音声認識部3-1において、統計階層データベース7-1の4つのすべての階層に対して音声認識処理プロセスが同時に走る始める。

音声認識結果出力部3-2は、最上位の60名のリストからなる最上位の統計階層データベース内の主催者名を、音声認識部3-1の音声認識結果を基に、音声認識結果の検索キー認識尤度の高い順に整列することにより音声認識結果テーブル（図7参照）を作成し、対話制御部4へ送る。

#### 【0038】

結果調整部4-1は、図7に示した音声認識結果テーブル中の音声検索キー候補に対して、あらかじめ定めた尤度しきい値3500を越える検索キー認識尤度を持つ音声検索キー有力候補として選択する。図7において、音声検索キー有力候補は「ジャパニーズJr」「キンキンクッツ」「パフェー」「ギリシア歌劇団」「チェコバレエ団」の5候補存在することがわかる。

#### 【0039】

そこで対話誘導部4-2は、音声検索キー有力候補が所定数2候補以上存在するので、次対話誘導条件の1を満たすことを把握し、関連属性情報がこの時点で

1つも得られていないことから、関連属性情報獲得のために検索キー確定関連質問を行うことを決定する。図8に示したように、検索キー属性データベース7-2には、統計階層データベース7-1の全データ（コンサート主催者350名）に対して、各データが有する属性項目の属性値として、それぞれのコンサートの主催月日（3月と4月の61日分のコンサートのみ対象とする）、主催曜日、主催会場（全国40ヶ所を対象とする）、コンサート主催都道府県、コンサート種別、コンサート主催者所属事務所等が付与してある。

【0040】

対話誘導部4-2は、図8に示した検索キー属性データベースを基に、コンサート主催月日を検索キー確定関連質問として尋ねることを決定し、応答文／質問文作成部4-2に対して「コンサートの主催月日はいつですか」という検索キー確定関連質問を、応答文／質問文作成部4-2に作成するよう指示する。

音声出力部6は音声出力装置9を用いてコンサート主催日を尋ねる検索キー確定関連質問を、ユーザに提示する。

ユーザからの上記検索キー確定関連質問への応答「3月3日」が、再び音声入力部2から入力される。

【0041】

音声識別部3に送られた、ユーザの応答「3月3日」に対して、音声認識部3-1は音声認識装置8を用いて音声認識処理を行い、音声認識結果出力部3-2は、検索キー属性データベース7-2の主催月日欄を照合することにより、コンサート主催月日候補に対して認識尤度の大きい順に並べた関連情報認識結果テーブル（図9参照）を作成し対話制御部4へ送る。

【0042】

この時点で、第二統計階層データベース（主催者150名含有）に対する音声認識及び音声認識結果テーブル作成が終了している。第二統計階層データベースに対する音声認識結果テーブルを図10に示した。

対話制御部4の結果調整部4-1は、第二統計階層データベースを参照し、コンサート主催日に関する関連情報認識結果テーブル（図10）内の主催月日候補と、第二統計階層データベースに対する音声認識結果テーブル中の各音声検索キ



一候補との関連性が認められた音声検索キー候補に対して、関連性の有る音声検索キー候補の検索キー認識尤度と関連属性情報候補の関連情報認識尤度を正規化し乗算することを音声検索キー関連性計算部 5 へ指示する。

【0043】

音声検索キー関連性計算部 5 はまず始めに、図 10 の音声認識結果テーブルにおける検索キー認識尤度を図 10 の最右欄のように正規化する。そして、図 10 の音声認識結果テーブルの尤度しきい値 3500 を超える 5 つの音声検索キー候補に対して、検索キー補助情報データベース 7-2 を用いて、「グリコ裕三／3 月 30 日」「イブニング娘／4 月 10 日」「ジャパニーズ Jr／5 月 30 日」「group／3 月 3 日」「キンキンクッツ／4 月 10 日」というコンサート主催日の情報を、関連属性情報として取得する。そして、図 9 に示した関連情報認識結果テーブル内の主催日に対しても関連情報認識尤度を最右欄のように正規化し、第二統計階層データベースから得られた音声認識結果テーブル（図 10）中の尤度しきい値 3500 を越える 5 つの音声検索キー「グリコ裕三」「イブニング娘」「ジャパニーズ Jr」「group」「キンキンクッツ」のコンサート主催日に該当する場合は、関連情報認識結果テーブル内の正規化した関連情報認識尤度を、音声認識結果テーブル内の各音声検索キー候補の正規化した検索キー認識尤度とかけあわせ、「グリコ裕三」「イブニング娘」「ジャパニーズ Jr」「group」「キンキンクッツ」の新たな認識尤度とする。

【0044】

つまり、音声検索キー候補である音声認識結果テーブル（図 10）内のグリコ裕三と、関連情報認識結果テーブル内のコンサート主催日の関連性、同様に、「イブニング娘」や「ジャパニーズ Jr」と関連情報認識結果テーブル内のコンサート主催日の関連性を、音声検索キー候補である各コンサート主催者の正規化した検索キー認識尤度と、各コンサートの主催日に対する認識結果を正規化した関連情報認識尤度の積を計算することにより、新たな認識尤度とする。

【0045】

この場合、図 11 に示すように、「グリコ裕三」のコンサート主催日である「3 月 30 日」の関連情報認識結果テーブル内の正規化した関連情報認識尤度は 0

．0055であることから、「グリコ裕三」の新たな認識尤度は $0.0080 \times 0.0055 = 0.00044$ 、同様に「イブニング娘」の認識正規化尤度 $0.0077$ に対する新たな認識尤度は「4月10日」の正規化した関連情報認識尤度 $0.0080$ との積をとり $0.00062$ 、「ジャパニーズJr」の新たな認識尤度は「5月30日」の正規化した関連情報認識尤度との積をとるのだが、「5月30日」という主催日は認識対象語彙データではないことから、認識不可能であるため関連情報認識結果が得られていないので、「ジャパニーズJr」の新たな認識尤度は0とする。「group」の新たな認識尤度は「3月3日」の正規化した関連情報認識尤度 $0.0077$ との積をとり $0.00054$ 、そして、「キンキンクッツ」の新たな認識尤度は「4月10日」の正規化した関連情報認識尤度 $0.0080$ との積をとり $0.00054$ となる。ここで認識尤度を統合した結果を正規化すると図11の下段表最右欄のようになる。

【0046】

結果調整部4-1は、上記のようにして定めた第二統計階層データベースの音声検索キー候補に対して正規化及び統合して新たな認識尤度を計算した結果を対話誘導部4-2に送る。

対話誘導部4-2は、正規化した統合認識尤度を見て、第二統計階層データベースに対する検索キー認識尤度の尤度しきい値を $0.2590$ と定める。この尤度しきい値の決定は、例えば最大尤度値より所定値以下の値とする。そして、尤度しきい値 $0.2590$ を越える音声検索キー候補の数に応じた対話誘導を始める。ここでは認識尤度が $0.2590$ を超える音声検索キー有力候補が、「イブニング娘」と「group」の2候補存在することが図11から分かる。

【0047】

そこで、対話誘導部4-2は正規化統合認識結果テーブル中の音声検索キー有力候補が所定数2以下であることから、検索キー属性データベース7-2を参照して、新たな関連属性情報を獲得し有力候補からの絞り込みを行うことを決定し、新たな関連属性情報として、コンサートの主催会場を尋ねることを決定する。

応答文／質問文作成部4-3は、コンサート主催会場を尋ねる「コンサートの主催会場をお答え下さい」という検索キー確定関連質問を作成し、音声出力部6

から検索キー確定関連質問が出力される。

【0048】

ユーザからの「東京ドーム」という応答音声が入力部2から入力され、音声識別部3へと送られる。音声識別部3の音声認識部3-1においてコンサート主催会場候補の音声認識処理が行われ、音声認識結果出力部3-2において各候補の関連情報認識尤度が計算され、関連情報認識結果テーブルが対話制御部4へ送られる。図12に、関連属性情報として得られたコンサート主催会場に対する関連情報認識結果テーブルを示す。図12の関連情報認識結果テーブル内の最右欄は正規化した認識尤度である。

【0049】

そして、結果調整部4-1は、今得られたコンサート主催会場だけでなく、関連情報認識結果テーブル保存場所に保存されている先ほどのコンサート主催月日を尋ねることにより得られた主催月日情報の両関連属性情報と、今絞込みの対象となっている第二統計階層データベース中の音声検索キー有力候補の関連性を判断して認識尤度の統合を音声検索キー関連性計算部5に指示する。

【0050】

音声検索キー関連性計算部5は、検索キー属性データベース7-2を参照して「イブニング娘」「group」の両音声検索キー有力候補と、コンサート主催日認識結果と新たに得られたコンサート主催会場認識結果の両関連属性情報候補との関連性が認められた場合、検索キー認識尤度と、各関連属性情報の関連情報認識尤度の統合を行う。

【0051】

すなわち実施例の場合、コンサート主催会場として、図12に示したように「東京ドーム」「横浜アリーナ」「福岡ドーム」「大阪城ホール」「NHKホール」などが関連属性情報候補として得られる。図11の下段表最右欄の「イブニング娘」「group」の正規化新尤度は、「イブニング娘」「group」の両音声検索キー有力候補の検索キー認識尤度と、コンサート主催月日情報の関連情報認識尤度を正規化及び統合した値であることから、新たに得られたコンサート主催会場候補の関連情報認識尤度を図11の下段最右欄の値と統合することで、

コンサート主催月日情報と主催会場情報の2つの関連属性情報との統合を行ったことになる。音声検索キー有力候補「イブニング娘」「group」のコンサート主催会場(図12)の関連性を、検索キー属性データベース7-2から判定する。

#### 【0052】

その結果、図13に示すように、「イブニング娘」は「名古屋ドーム」という関連属性を持つことから、それぞれの正規化した認識尤度0.2897、0.0397を乗算し「イブニング娘」の新たな認識尤度は0.01150、「group」は「東京ドーム」という関連属性を持つことから、正規化した認識尤度を乗算して、 $0.2593 \times 0.0833$ で「group」の新たな認識尤度は0.02160(図13参照)となる。

#### 【0053】

上記結果から、対話誘導部4-2は対話誘導方式に従い、認識尤度の統合の結果最も大きな検索キー認識尤度を算出した音声検索キー有力候補「group」を音声検索キーと同定し、応答文/質問文作成部4-3にユーザへの提示文作成を指示する。

音声出力部6は「groupの3月3日、東京ドームコンサートですね」という同定候補提示文を出力する。

#### 【0054】

ユーザのこの提示に対する「YES」という応答が、音声入力部2から入力され、音声識別部3に送られて、音声認識部3-1はYes/Noデータベース7-4を用いて音声認識処理をし、音声認識結果出力部3-2は認識結果を対話制御部4へ送る。

結果調整部は、音声認識結果出力部3-2から送られてきた「Yes」という認識結果を対話誘導部4-2へ送り、対話誘導部4-2は正しい音声検索キーが同定できたことを判断し対話終了を決定する。

#### 【0055】

この発明の装置の各機能構成はコンピュータによりプログラムを解読実行させることにより機能させることもできる。

【0056】

## 【発明の効果】

以上の説明からわかるように、音声認識対象語彙数が大規模な場合、現在の音声認識技術では認識処理に時間がかかり、さらに認識精度が100%でないことから、ユーザから音声により要求されるタスクを与えられた一定時間で達成することは困難である。すなわち、ユーザはシステムが音声認識を行っている間、待機しなければならない、待機した上で提示された候補が誤認識の場合は、正しい候補が提示されるまで質疑応答を繰り返したり、再び音声入力を指示され待機することになり、対オペレータとの対話のような自然な対話の中でタスクを達成することは困難である。

【0057】

この発明によれば、音声認識用データベースを、アクセス頻度などの統計的情報に基づいてすべてのデータに重要度を設け、重要度に基づいてデータを細分化及び階層構造化した複数の統計階層データベースとして保持し、各階層データベース含有データ数の相違による認識時間の差を利用して音声認識処理の仮想的な実時間性を実現し、さらに音声認識処理の認識尤度にしきい値を設けることで、信頼性の高い認識結果が少数の場合は、関連属性情報を尋ねることで効果的な絞り込みを行い、信頼性の高い認識結果が所定数よりも多い場合、或いは、信頼性の高い認識結果が存在しない場合、或いは、第一の候補が該当検索キーではないとユーザから否定された場合などは、最上位の階層に該当検索キー候補が含まれていない可能性があることから、下位の統計階層データベースに認識対象を移動し、関連属性情報との統合を行うことで音声認識率の欠陥を補い、さらには検索キー確定関連質問で対話をつなぐことにより、あたかも全データに対して音声認識処理を行っているかのように見せ、自然な対話を実現する。

【0058】

尚、オペレータ対応でタスク達成を行っている、座席の価格を属性として座席名を同定する座席同定や、路線名を属性として駅名を同定する駅名検索に対しても、この発明は容易に拡張可能であると考えられる。

さらに、人名検索などに関しては、住所、性別、職業、年齢、電話番号などの

複数の属性をあらかじめ関連属性情報として保持し、統合して利用することで、検索キーである人名の同定が可能になると考えられる。

【図面の簡単な説明】

【図 1】

この発明によるデータベース検索装置の実施形態の機能構成を示すブロック図。

【図 2】

この発明で利用する統計階層データベースの一例を示す図。

【図 3】

音声検索キー候補に対する認識尤度を算出した音声認識結果テーブルの一例を示す図。

【図 4】

この発明で利用する検索キー属性データベースの一例を示す図。

【図 5】

検索キー確定関連質問に対するユーザの応答に対する音声認識結果を示す関連情報認識結果テーブルの一例を示す図。

【図 6】

コンサートチケット予約システムにおける音声認識用統計階層データベースの一例を示す図。

【図 7】

コンサートチケット予約システムにおける第一統計階層データベースに対する音声認識結果テーブルの一例を示す図。

【図 8】

コンサートチケット予約システムにおける検索キー属性データベースの一例を示す図。

【図 9】

コンサート主催月日を尋ねる検索キー確定関連質問に対する応答から得られる関連情報認識結果テーブルの一例を示す図。

【図 10】

コンサートチケット予約システムにおける第二統計階層データベースに対する音声認識結果テーブルの一例を示す図。

【図 1 1】

第二統計階層データベースと主催月日の関連情報認識結果テーブルとの統合例を示す図。

【図 1 2】

コンサート主催会場を尋ねる検索キー確定関連質問に対する応答から得られる関連情報認識結果テーブルの一例を示す図。

【図 1 3】

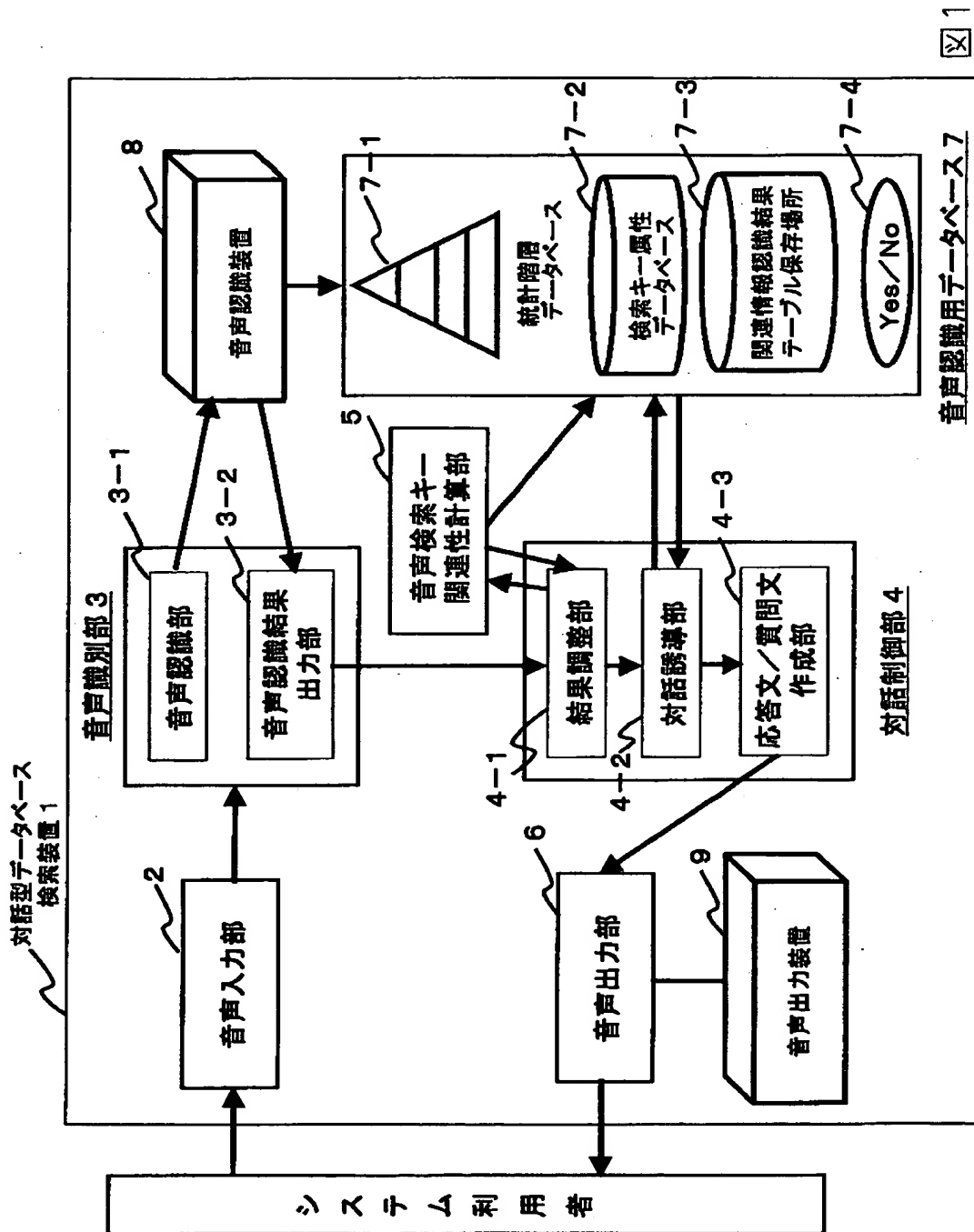
第二統計階層データベース中の音声検索キー有力候補と主催月日、及び主催会場情報の統合例を示す図。

【図 1 4】

この発明における対話制御部 4（図 1）の処理手順の例を示す流れ図。

【書類名】 図面

【図 1】





【図 2】

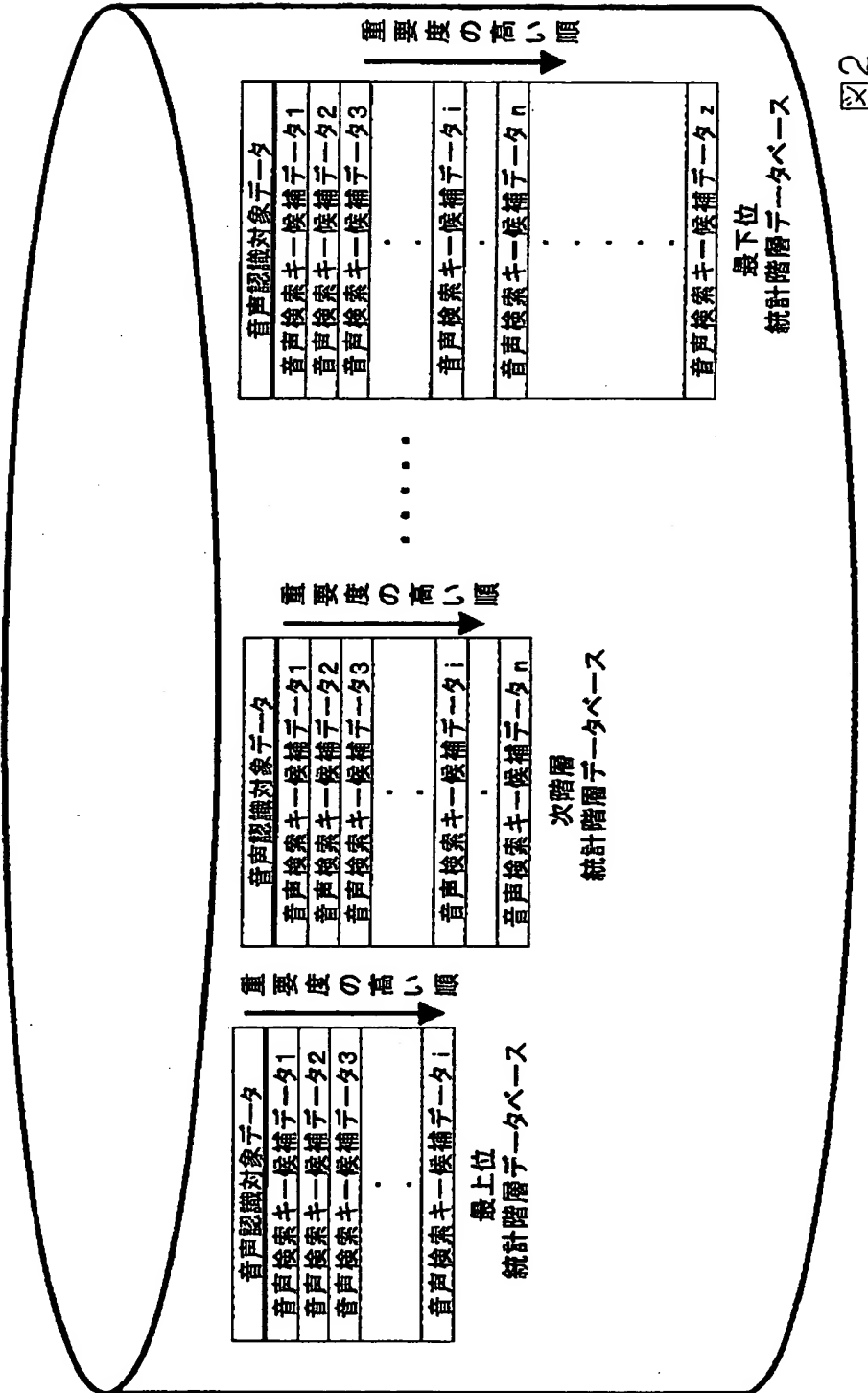


図 2

【図 3】

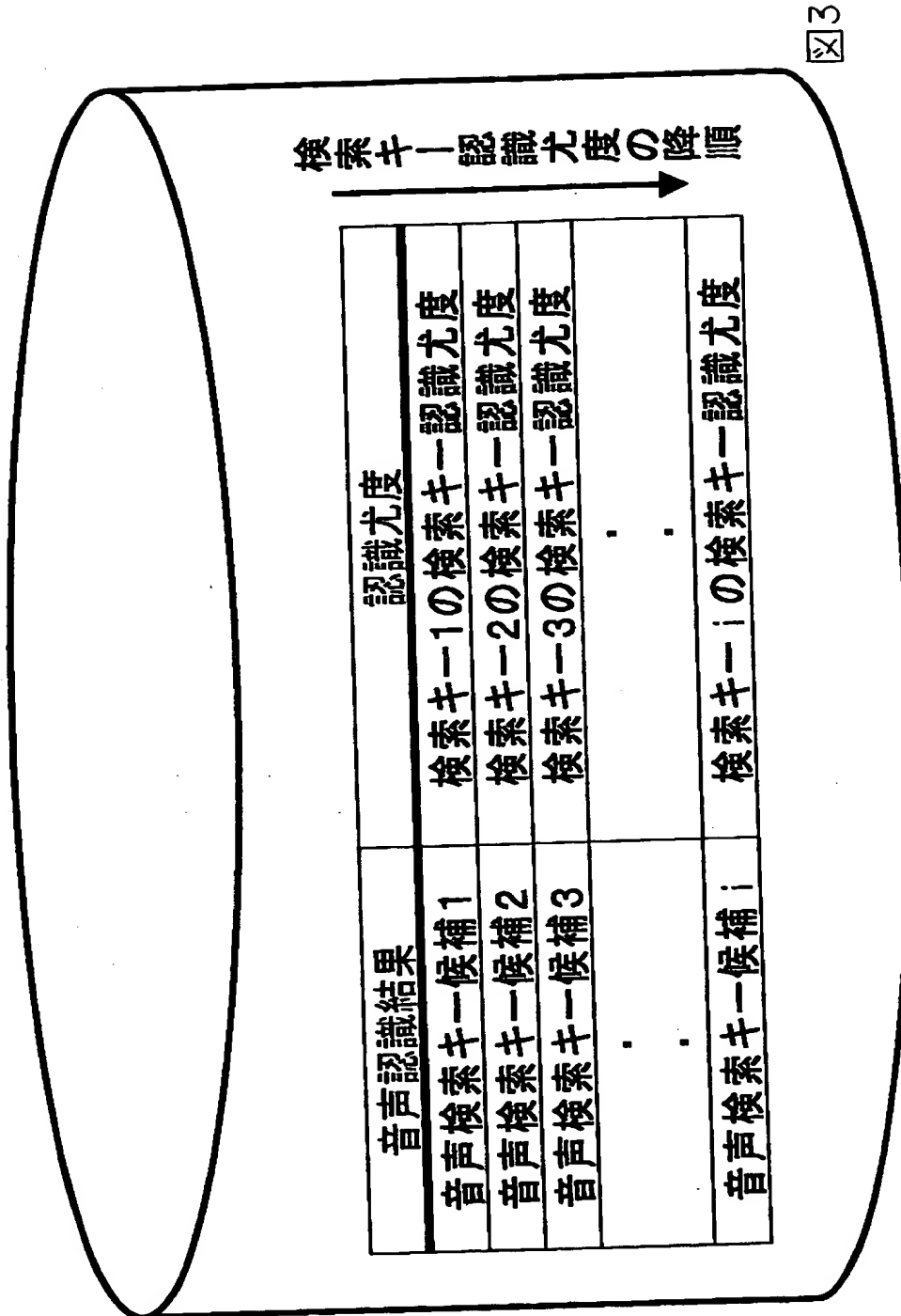


図 3

【図 4】

|         | 関連属性 1     | . | 関連属性 m     |
|---------|------------|---|------------|
| 音声検索キ-1 | 検索キ-1の関連属性 | . | 検索キ-1の関連属性 |
| 音声検索キ-2 | 検索キ-2の関連属性 | . | 検索キ-2の関連属性 |
| 音声検索キ-3 | 検索キ-3の関連属性 | . | 検索キ-3の関連属性 |
| .       | .          | . | .          |
| 音声検索キ-i | 検索キ-iの関連属性 | . | 検索キ-iの関連属性 |
| .       | .          | . | .          |
| .       | .          | . | .          |
| 音声検索キ-n | 検索キ-nの関連属性 | . | 検索キ-nの関連属性 |
| .       | .          | . | .          |
| .       | .          | . | .          |
| .       | .          | . | .          |
| 音声検索キ-z | 検索キ-zの関連属性 | . | 検索キ-zの関連属性 |

図 4

【図 5】

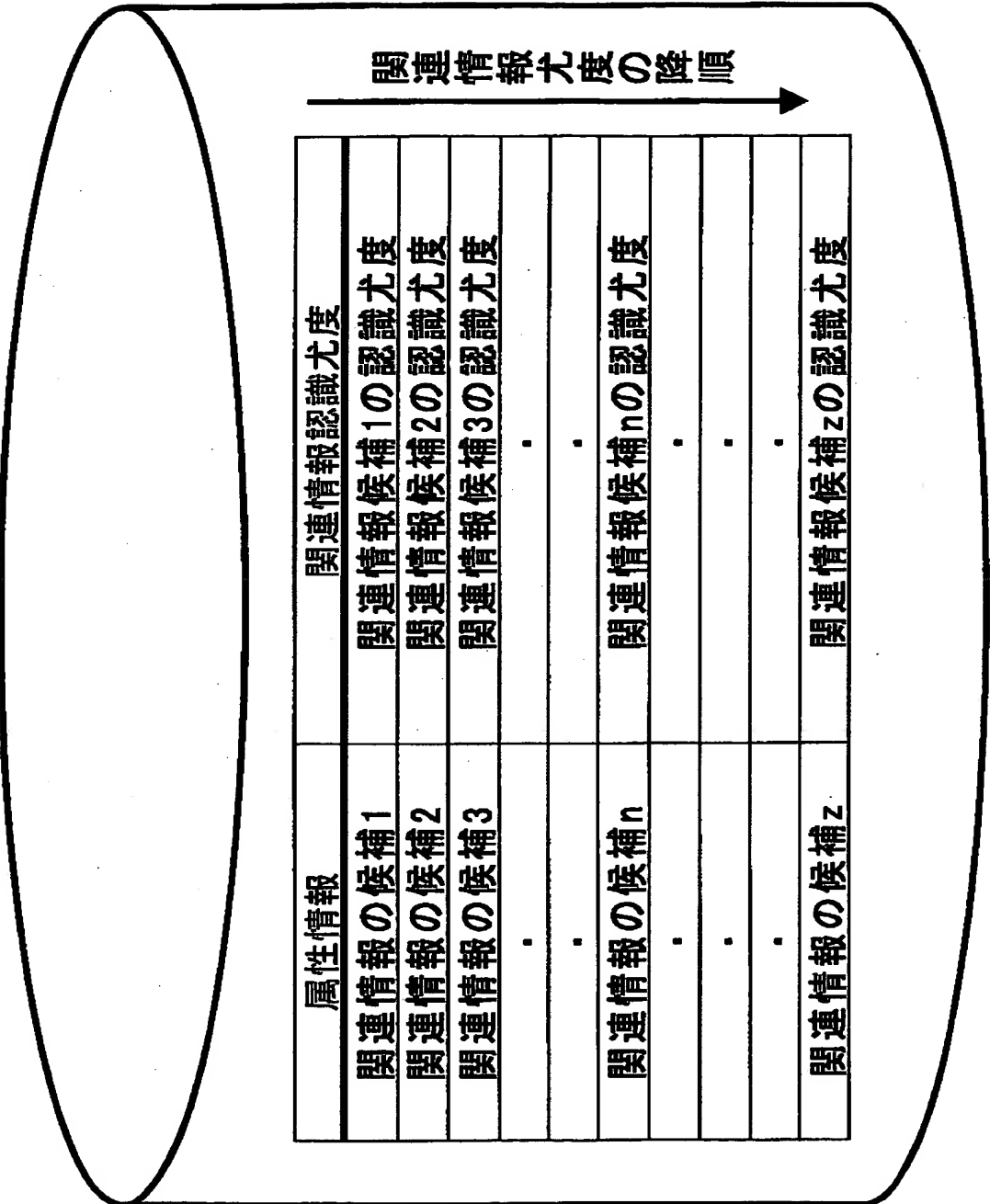


図5

【図6】

|     | 第一階層     | 第二階層     | 第三階層     | 第四階層     |
|-----|----------|----------|----------|----------|
| 1   | 奈室安美恵    | 奈室安美恵    | 奈室安美恵    | 奈室安美恵    |
| 2   | SNAP     | SNAP     | SNAP     | SNAP     |
| 3   | パフェ-     | パフェ-     | パフェ-     | パフェ-     |
| 4   | ジャパニーズJr | ジャパニーズJr | ジャパニーズJr | ジャパニーズJr |
| 5   | SILVER   | SILVER   | SILVER   | SILVER   |
| 6   | キンキンクッツ  | キンキンクッツ  | キンキンクッツ  | キンキンクッツ  |
| .   | .        | .        | .        | .        |
| .   | .        | .        | .        | .        |
| .   | .        | .        | .        | .        |
| 59  | ギリシア歌劇団  | ギリシア歌劇団  | ギリシア歌劇団  | ギリシア歌劇団  |
| 60  | チェコバレエ団  | チェコバレエ団  | チェコバレエ団  | チェコバレエ団  |
| .   |          | .        | .        | .        |
| .   |          | .        | .        | .        |
| .   |          | .        | .        | .        |
| 90  |          | group    | group    | group    |
| .   |          | .        | .        | .        |
| 149 |          | グリコ裕三    | グリコ裕三    | グリコ裕三    |
| 150 |          | イブニング娘   | イブニング娘   | イブニング娘   |
| .   |          | .        | .        | .        |
| .   |          | .        | .        | .        |
| .   |          | .        | .        | .        |
| .   |          | .        | .        | .        |
| .   |          | .        | .        | .        |
| 249 |          | 中年隊      | 中年隊      | 中年隊      |
| 250 |          | 中海美穂     | 中海美穂     | 中海美穂     |
| .   |          |          |          | .        |
| .   |          |          |          | .        |
| .   |          |          |          | .        |
| .   |          |          |          | .        |
| 350 |          |          |          | 吉木歌劇団    |

図6

【図 7】

検索キー認識尤度の降順

→

|    | 音声検索キー候補 | 認識尤度     | 正規化尤度  |
|----|----------|----------|--------|
| 1  | ジャパニーズJr | 4,500    | 0.0375 |
| 2  | キンキンクッツ  | 4,300    | 0.0358 |
| 3  | パフエ-     | 3,900    | 0.0325 |
| 4  | ギリシア歌劇団  | 3,700    | 0.0308 |
| 5  | チエコバレエ団  | 3,700    | 0.0308 |
| 6  | SILVER   | 3,400    | 0.0283 |
| .  | .        | .        | .      |
| .  | .        | .        | .      |
| 60 | SPED     | 1,600    | 0.0133 |
|    |          | 計120,000 | 1.0000 |

図7

【図 8】

|     | 検索キー候補   | 主催月日  | 主催会場   | 曜日 | 都道府県 | 種別      |
|-----|----------|-------|--------|----|------|---------|
| 1   | 奈室安美恵    | 4月12日 | 渋谷公会堂  | 月  | 東京都  | 歌謡      |
| 2   | SNAP     | 5月30日 | 大阪城ホール | 日  | 大阪府  | 歌謡      |
| 3   | パフエ-     | 4月3日  | 渋谷公会堂  | 土  | 東京都  | 歌謡      |
| 4   | ジャパニーズJr | 5月30日 | NHKホール | 日  | 東京都  | 歌謡      |
| 5   | SILVER   | 4月20日 | 名古屋ドーム | 火  | 愛知県  | 歌謡      |
| 6   | キンキンクッツ  | 4月10日 | 渋谷公会堂  | 土  | 東京都  | 歌謡      |
| .   | .        | .     | .      | .  | .    | .       |
| .   | .        | .     | .      | .  | .    | .       |
| 59  | ギリシア歌劇団  | 4月20日 | 国際ホール  | 火  | 東京都  | 歌劇      |
| 60  | チエコバレエ団  | 4月21日 | 国際ホール  | 水  | 東京都  | 芸術      |
| .   | .        | .     | .      | .  | .    | .       |
| .   | .        | .     | .      | .  | .    | .       |
| 90  | group    | 3月3日  | 東京ドーム  | 水  | 東京都  | 歌謡      |
| .   | .        | .     | .      | .  | .    | .       |
| 149 | グリコ裕三    | 3月30日 | 大阪城ホール | 火  | 大阪府  | ディナーショー |
| 150 | イブニング娘   | 4月10日 | 名古屋ドーム | 土  | 愛知県  | 歌謡      |
| .   | .        | .     | .      | .  | .    | .       |
| .   | .        | .     | .      | .  | .    | .       |
| .   | .        | .     | .      | .  | .    | .       |
| 249 | 中年隊      | 4月2日  | 厚生年金会館 | 金  | 東京都  | 歌謡      |
| 250 | 中海美穂     | 4月21日 | 福岡市民会館 | 水  | 福岡県  | 歌謡      |
| .   | .        | .     | .      | .  | .    | .       |
| .   | .        | .     | .      | .  | .    | .       |
| 350 | 吉木歌劇団    | 3月3日  | 大阪府立会館 | 水  | 大阪府  | 漫才      |

図8

【図 9】

|    | 属性情報候補 | 関連情報尤度   | 正規化尤度  |
|----|--------|----------|--------|
| 1  | 3月5日   | 5,600    | 0.0102 |
| 2  | 3月10日  | 5,400    | 0.0098 |
| 3  | 3月4日   | 4,800    | 0.0090 |
| 4  | 4月10日  | 4,400    | 0.0080 |
| 5  | 3月3日   | 4,200    | 0.0077 |
| 6  | 4月3日   | 3,500    | 0.0064 |
| 7  | 3月30日  | 3,000    | 0.0055 |
| .  | .      | .        | .      |
| .  | .      | .        | .      |
| 61 | 4月20日  | 1,100    | 0.0020 |
|    |        | 計549,000 | 1.0000 |

関連情報尤度の降順



図 9



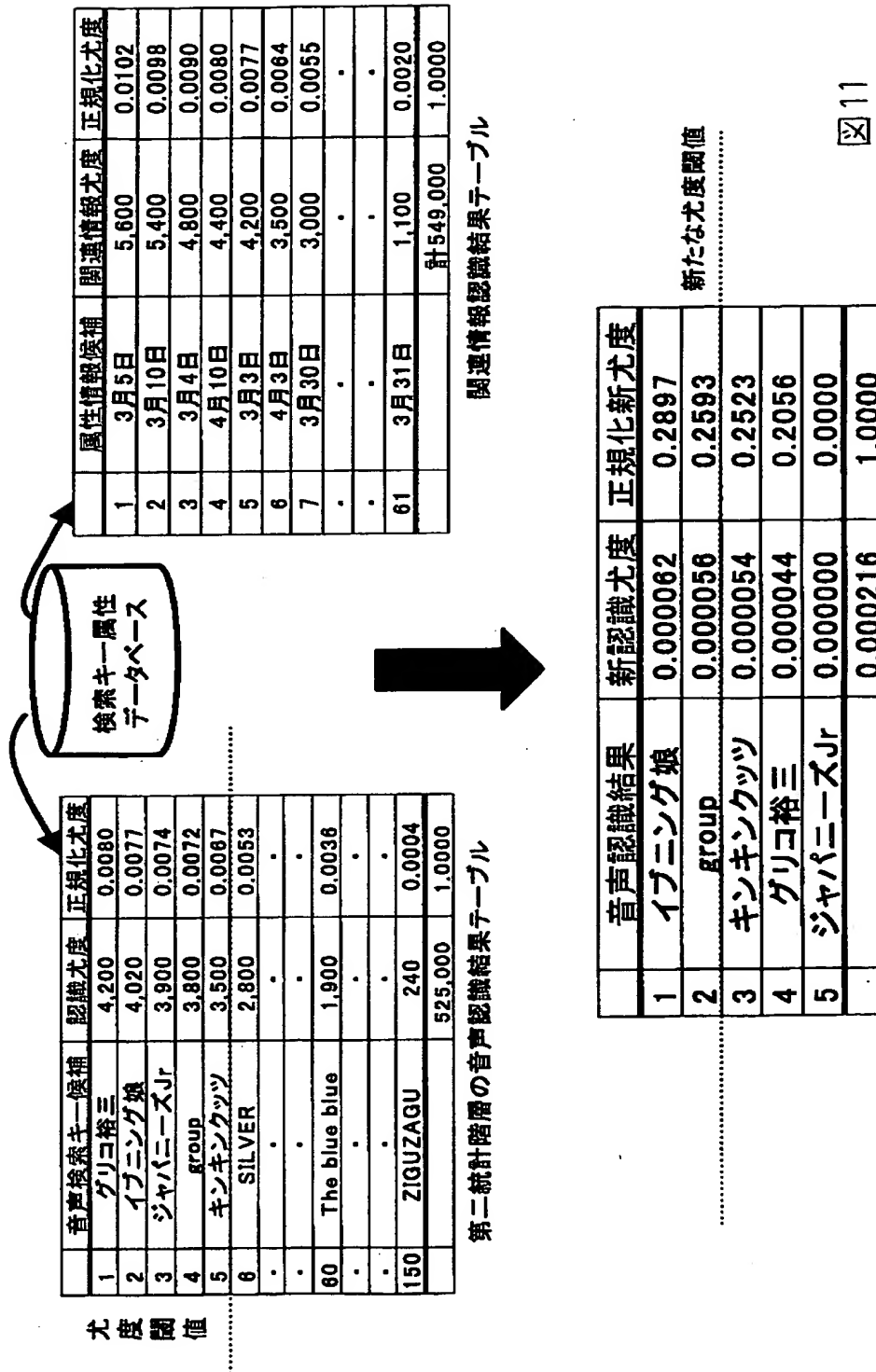
【図 1 0】

|     | 音声検索キー候補      | 認識尤度    | 正規化尤度  |
|-----|---------------|---------|--------|
| 1   | グリコ裕三         | 4,200   | 0.0080 |
| 2   | イブニング娘        | 4,020   | 0.0077 |
| 3   | ジャパニーズJr      | 3,900   | 0.0074 |
| 4   | group         | 3,800   | 0.0072 |
| 5   | キンキンクッツ       | 3,500   | 0.0067 |
| 6   | SILVER        | 2,800   | 0.0053 |
| .   | .             | .       | .      |
| .   | .             | .       | .      |
| 60  | The blue blue | 1,900   | 0.0036 |
| .   | .             | .       | .      |
| .   | .             | .       | .      |
| 150 | ZIGUZAGU      | 240     | 0.0004 |
|     |               | 525,000 | 1.0000 |

検索キー認識尤度の降順

図10

【図11】



【図 12】

|    | 属性情報候補 | 関連情報尤度 | 正規化尤度  |
|----|--------|--------|--------|
| 1  | 東京ドーム  | 4,200  | 0.0833 |
| 2  | 横浜アリーナ | 3,900  | 0.0774 |
| 3  | 福岡ドーム  | 3,700  | 0.0734 |
| 4  | 大阪城ホール | 2,200  | 0.0436 |
| 5  | NHKホール | 2,000  | 0.0397 |
| 6  | 名古屋ドーム | 2,000  | 0.0397 |
| .  | .      | .      | .      |
| .  | .      | .      | .      |
| 40 | 苗場プリンス | 1,400  | 0.0278 |
|    |        | 50,400 | 1.0000 |

関連情報尤度の降順

図 12

【図 13】

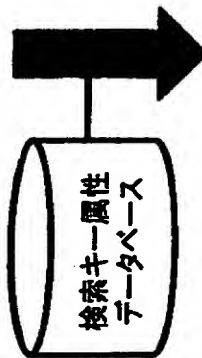
|    | 属性情報候補 | 関連情報尤度 | 正規化尤度  |
|----|--------|--------|--------|
| 1  | 東京ドーム  | 4,200  | 0.0833 |
| 2  | 横浜アリーナ | 3,900  | 0.0774 |
| 3  | 福岡ドーム  | 3,700  | 0.0734 |
| 4  | 大阪城ホール | 2,200  | 0.0436 |
| 5  | NHKホール | 2,000  | 0.0397 |
| 6  | 名古屋ドーム | 2,000  | 0.0397 |
| .  | .      | .      | .      |
| .  | .      | .      | .      |
| 40 | 苗場プリンス | 1,400  | 0.0278 |
|    |        | 50,400 | 1.0000 |

主催会場の関連情報認識結果テーブル

尤度 閾 値

|   | 音声認識結果   | 新認識尤度    | 正規化新尤度 |
|---|----------|----------|--------|
| 1 | イブニング娘   | 0.000062 | 0.2897 |
| 2 | group    | 0.000056 | 0.2593 |
| 3 | キンキンクッツ  | 0.000054 | 0.2523 |
| 4 | グリコ裕三    | 0.000044 | 0.2056 |
| 5 | ジャパニーズJr | 0.000000 | 0.0000 |
|   |          | 0.000216 | 1.0000 |

第二統計階層の音声認識結果とコンサート主催月日の統合結果



|   | 音声認識結果 | 新認識尤度   | 正規化新尤度 |
|---|--------|---------|--------|
| 1 | group  | 0.02160 | 0.6123 |
| 2 | イブニング娘 | 0.01150 | 0.3260 |

図 13

【図 14】

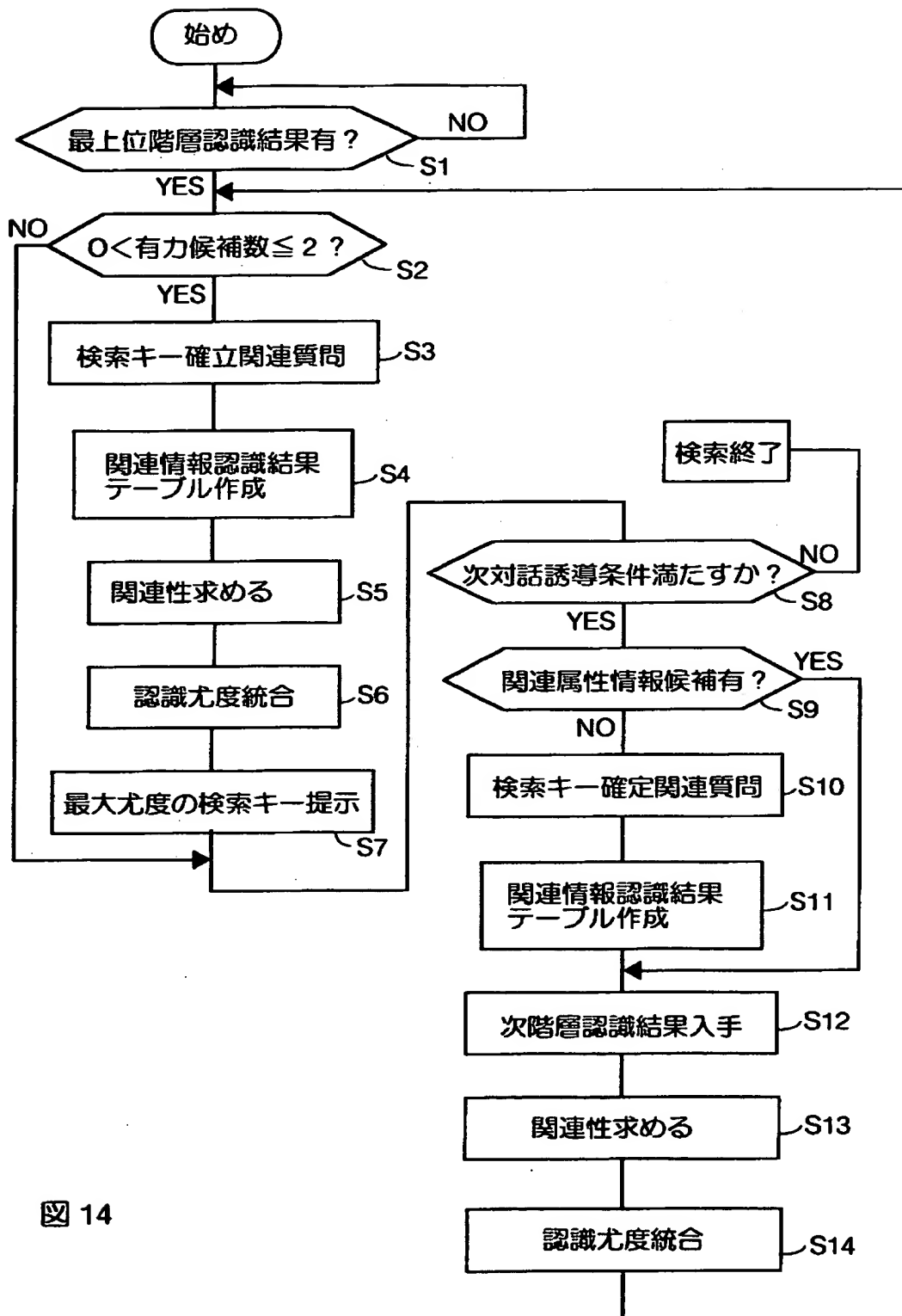


図 14

【書類名】 要約書

【要約】

【課題】 ユーザに待機時間を感じさせない。

【解決手段】 音声認識用データベースを、重要なデータから順に所定数のデータを選択して部分データベースとし、下位階層ほど含有データ数を多くし、最下位は全データを含むように階層化し、認識対象音声検索キーに対し、各階層データベースごとに並列に音声認識を行い、最上位階層についての認識結果中の有力候補が2以下で0でない場合は、検索キー確定関連質問を行い、その回答を音声認識処理して、関連属性情報候補を得、これと有力候補との関連があるものについて、両者の尤度を統合して検索キー候補の尤度とし、その最大のものを提示し、同定できない場合は、次階層の認識結果について同様のことを行う。

【選択図】 図14

出 願 人 履 歴 情 報

識別番号 [000004226]

1. 変更年月日 1995年 9月21日  
[変更理由] 住所変更  
住 所 東京都新宿区西新宿三丁目19番2号  
氏 名 日本電信電話株式会社
2. 変更年月日 1999年 7月15日  
[変更理由] 住所変更  
住 所 東京都千代田区大手町二丁目3番1号  
氏 名 日本電信電話株式会社